Title	LMI-based estimation of scene points in vision systems with generalized cameras
Author(s)	Chesi, G
Citation	IEEE Transactions on Automatic Control, 2014, v. 59 n. 11, p. 2996-3001
Issued Date	2014
URL	http://hdl.handle.net/10722/211751
Rights	Creative Commons: Attribution 3.0 Hong Kong License

LMI-Based Estimation of Scene Points in Vision Systems with Generalized Cameras

Graziano Chesi

***Abstract*—This paper considers the problem of estimating the position of a scene point from its image projections onto generalized cameras, i.e., cameras that can be modeled by a spherical projection followed by a perspective one. The sought estimate is defined through a geometric criterion, specifically the minimization of the angles between the projections on the sphere of the available image points and the corresponding projections of the estimate. A solution based on convex optimization with linear matrix inequalities (LMIs) is proposed for addressing this problem, which provides a candidate of the sought estimate. Moreover, a condition is provided for establishing exactness of the found candidate, i.e., establishing whether the found candidate is a minimizer of the considered geometric criterion.**

***Index Terms*—Generalized camera, linear matrix inequality (LMI), scene estimation, vision system.**

I. Introduction

It is well-known that estimating the position of a scene point from its image projections onto two or more cameras is a fundamental problem in vision systems, also known as multiple-view triangulation, see, e.g., [1]. Indeed, this problem has numerous applications, in particular an important one is in vision-based control of robots, see, e.g., [2], [13]. Due to image noise and calibration errors, only an estimate of the sought scene point can be obtained, whose accuracy depends on the criterion chosen to match the available image points with the image projections of the candidate estimate.

For perspective cameras, numerous contributions can be found, which typically consider a geometric criterion for defining the estimate of the sought scene point. A commonly adopted geometric criterion is the minimization of the reprojection error in the L_2 norm, for which several solutions have been proposed. In [3], the authors show how the exact solution with two views can be obtained by computing the roots of a one-variable polynomial of degree six. For the case of three views, the exact solution is obtained in [4] by solving a system of polynomial equations through methods from computational commutative algebra. For these cases and others with more views, [5] derives a solution based on convex optimization exploiting the fundamental matrix.

This paper considers multiple-view triangulation in a vision system with generalized cameras, i.e., cameras that can be modeled by a spherical projection followed by a perspective one. This class of cameras include the classic perspective cameras as well as non-perspective ones such as fisheye cameras. The sought estimate is defined through a geometric criterion, specifically the minimization of the angles between the projections on the sphere of the available image points and the corresponding projections of the estimate. A solution based on convex optimization with linear matrix inequalities (LMIs) is proposed for addressing this problem, which provides a candidate of the sought estimate. Moreover, a condition is provided for establishing exactness of the found candidate, i.e., establishing whether the found candidate is a minimizer of the considered geometric criterion. A conference version of this paper (without the exactness condition) appeared in [6].

Manuscript received September 23, 2012; revised June 26, 2013; accepted November 7, 2013. Date of publication August 28, 2014; date of current version October 21, 2014. This work was supported in part by the Research Grants Council of Hong Kong under Grant HKU711213E. Recommended by Associate Editor D. Regruto.

The author is with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong (e-mail: chesi@eee.hku.hk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2014.2351657

II. Preliminaries

Notation: $\mathbb{R}$: real numbers set; $\mathbb{P}^n$: set of vectors in $\mathbb{R}^n$ with last entry equal to 1; $\mathbf{x} \in \mathbb{R}^3$: scene point; $\mathbf{O}_i \in \mathbb{R}^{3\times 3}$: rotation matrix of the ith generalized camera; $\mathbf{c}_i \in \mathbb{P}^3$: center of the spherical projection; $\xi_i \in \mathbb{R}$: distance between $\mathbf{c}_i$ and the center of the perspective projection given by $\mathbf{c}_i - \xi_i \mathbf{O}_i \mathbf{e}_3$; $\mathbf{K}_i \in \mathbb{R}^{3\times 3}$: intrinsic parameters matrix; $\mathbf{a}_i \in \mathbb{R}^3$: spherical projection of $\mathbf{x}$; $\mathbf{b}_i \in \mathbb{P}^3$: image projection of $\mathbf{x}$ in normalized coordinates; $\mathbf{p}_i \in \mathbb{P}^3$: image projection of $\mathbf{x}$ in pixel coordinates; $\|\mathbf{v}\| \in \mathbb{R}$: Euclidean norm of $\mathbf{v}$; $\mathbf{v}^T$: transpose of $\mathbf{v}$; $\mathbf{e}_j \in \mathbb{R}^3$: jth column of the 3×3 identity matrix; SOS: sum of squares of polynomials; subject to (s.t.).

We consider n generalized cameras observing a scene point $\mathbf{x}$. The ith generalized camera consists of a spherical projection followed by a perspective one. Such a camera can be modeled using the so called unified model, see, e.g., [7]. In this model, the spherical projection of $\mathbf{x}$ is

$$\mathbf{a}_i = \boldsymbol{\sigma}_i(\mathbf{x}) \tag{1}$$

where

$$\boldsymbol{\sigma}_i(\mathbf{x}) = \frac{\mathbf{O}_i^T(\mathbf{x} - \mathbf{c}_i)}{\|\mathbf{O}_i^T(\mathbf{x} - \mathbf{c}_i)\|}. \tag{2}$$

The image projection of $\mathbf{x}$ in normalized coordinates is

$$\mathbf{b}_i = \boldsymbol{\tau}_i(\mathbf{x}) \tag{3}$$

where

$$\boldsymbol{\tau}_i(\mathbf{x}) = \frac{1}{\mathbf{e}_3^T\mathbf{a}_i + \xi_i\|\mathbf{a}_i\|} \begin{pmatrix} \mathbf{e}_1^T\mathbf{a}_i \\ \mathbf{e}_2^T\mathbf{a}_i \\ \mathbf{e}_3^T\mathbf{a}_i + \xi_i\|\mathbf{a}_i\| \end{pmatrix}. \tag{4}$$

The image projection of $\mathbf{x}$ in pixel coordinates is

$$\mathbf{p}_i = \mathbf{K}_i\mathbf{b}_i. \tag{5}$$

The dependence of $\mathbf{p}_i$ on $\mathbf{x}$ is denoted as

$$\mathbf{p}_i = \boldsymbol{\phi}_i(\mathbf{x}). \tag{6}$$

Let $\hat{\mathbf{p}}_i$ and $\hat{\boldsymbol{\phi}}_i$ be the available estimates of $\mathbf{p}_i$ and $\boldsymbol{\phi}_i$ (the latter defined by the available estimates $\hat{\mathbf{K}}_i, \hat{\xi}_i, \hat{\mathbf{O}}_i, \hat{\mathbf{c}}_i$ of $\mathbf{K}_i, \xi_i, \mathbf{O}_i, \mathbf{c}_i$). The multiple-view triangulation problem for generalized cameras is

$$\text{given } \{(\hat{\mathbf{p}}_i, \boldsymbol{\phi}_i), i = 1, \ldots, n\}, \text{ estimate } \mathbf{x}. \tag{7}$$

III. Estimation

In order to address the multiple-view triangulation problem for generalized cameras, we introduce a geometric criterion which consists of determining the scene point that minimizes the angles between the projections on the sphere of the available image points and of the sought point itself. In particular, we define this criterion according to

$$\mathbf{x}^* = \arg\max_{\mathbf{x}} g(\mathbf{x}) \quad \text{s.t.} \quad \begin{cases} \mathbf{e}_3^T \hat{\mathbf{O}}_i^T (\mathbf{x} - \hat{\mathbf{c}}_i) \geq 0 \; \forall i = 1, \ldots, n \\ \|\mathbf{x} - \hat{\mathbf{c}}_i\| \geq \varepsilon \; \forall i = 1, \ldots, n \\ \rho \geq \|\mathbf{x}\| \end{cases} \tag{8}$$

where the first constraint ensures that $\mathbf{x}$ lies in front of the generalized cameras, the second ensures that $\mathbf{x}$ does not coincide with the centers $\hat{\mathbf{c}}_i$ (ε is an arbitrarily small positive number), and the third ensures that $\mathbf{x}$ is finite (ρ is an arbitrarily large positive number). The cost function $g(\mathbf{x})$ is

$$g(\mathbf{x}) = \frac{1}{n} \sum_{i=1,\ldots,n} \hat{\boldsymbol{\sigma}}_i(\mathbf{x})^T \hat{\mathbf{a}}_i \tag{9}$$

where $\hat{\boldsymbol{\sigma}}_i(\mathbf{x})$ is the estimated spherical projection of $\mathbf{x}$ and $\hat{\mathbf{a}}_i$ is the back projection of $\hat{\mathbf{p}}_i$ on the sphere of the ith generalized camera. In fact, since $\hat{\boldsymbol{\sigma}}_i(\mathbf{x})$ and $\hat{\mathbf{a}}_i$ are unitary norm vectors, it follows that

$$\hat{\boldsymbol{\sigma}}_i(\mathbf{x})^T \hat{\mathbf{a}}_i = \cos\theta_i \tag{10}$$

where θ_i is the angle between $\hat{\boldsymbol{\sigma}}_i(\mathbf{x})$ and $\hat{\mathbf{a}}_i$. The expression for $\hat{\boldsymbol{\sigma}}_i(\mathbf{x})$ is given by (2) replacing $\mathbf{O}_i$ and $\mathbf{c}_i$ with their available estimates, i.e.,

$$\hat{\boldsymbol{\sigma}}_i(\mathbf{x}) = \frac{\hat{\mathbf{O}}_i^T (\mathbf{x} - \hat{\mathbf{c}}_i)}{\left\| \hat{\mathbf{O}}_i^T (\mathbf{x} - \hat{\mathbf{c}}_i) \right\|} \tag{11}$$

while the expression for $\hat{\mathbf{a}}_i$ is obtained inverting (5) and (4), hence obtaining

$$\hat{\mathbf{a}}_i = \begin{pmatrix} \hat{\mathbf{f}}_i \\ \sqrt{1 - \|\hat{\mathbf{f}}_i\|^2} \end{pmatrix} \tag{12}$$

where

$$\hat{\mathbf{f}}_i = \hat{\zeta}_i \begin{pmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \end{pmatrix} \hat{\mathbf{b}}_i \tag{13}$$

with

$$\hat{\zeta}_i = \frac{1 - \hat{\xi}_i^2}{\sqrt{\hat{\xi}_i^2 + (1 - \hat{\xi}_i^2)\|\hat{\mathbf{b}}_i\|^2} - \hat{\xi}_i} \tag{14}$$

and

$$\hat{\mathbf{b}}_i = \hat{\mathbf{K}}_i^{-1} \hat{\mathbf{p}}_i. \tag{15}$$

Let us observe that the operator $\arg\max$ in (8) is well-defined since the feasible set is compact. Since the maximizer in (8) could be nonunique, this operator can return a set of points in the general case, and $\mathbf{x}^*$ is defined as any one of such points. Let us also observe that (8) can present local maxima since the cost function $g(\mathbf{x})$ is nonlinear. In order to cope with this difficulty, we propose a reformulation of (8) into an optimization problem over polynomials as follows.

Let us start by defining the optimal cost of (8) as

$$g^* = \max_{\mathbf{x}} g(\mathbf{x}) \quad \text{s.t.} \quad \begin{cases} \mathbf{e}_3^T \hat{\mathbf{O}}_i^T (\mathbf{x} - \hat{\mathbf{c}}_i) \geq 0 \; \forall i = 1, \ldots, n \\ \|\mathbf{x} - \hat{\mathbf{c}}_i\| \geq \varepsilon \; \forall i = 1, \ldots, n \\ \rho \geq \|\mathbf{x}\|. \end{cases} \tag{16}$$

Let $y_1, \ldots, y_n \in \mathbb{R}$ be additional variables, and let us define the extended variable $\mathbf{z} \in \mathbb{R}^{n+3}$ as

$$\mathbf{z} = (\mathbf{x}^T, y_1, \ldots, y_n)^T. \tag{17}$$

From $\mathbf{z}$, we define

$$h(\mathbf{z}) = \frac{1}{n} \sum_{i=1,\ldots,n} y_i \left(\hat{\mathbf{O}}_i^T (\mathbf{x} - \hat{\mathbf{c}}_i) \right)^T \hat{\mathbf{a}}_i \tag{18}$$

which will be used as a new cost function. Moreover, we define

$$l_i(\mathbf{z}) = y_i^2 (\mathbf{x} - \hat{\mathbf{c}}_i)^T (\mathbf{x} - \hat{\mathbf{c}}_i) - 1 \tag{19}$$

for $i = 1, \ldots, n$, and

$$m_j(\mathbf{z}) = \begin{cases} y_i & \text{with } i = j \\ & \text{if } j \leq n \\ \mathbf{e}_3^T \hat{\mathbf{O}}_i^T (\mathbf{x} - \hat{\mathbf{c}}_i) & \text{with } i = j - n \\ & \text{if } n+1 \leq j \leq 2n \\ (\mathbf{x} - \hat{\mathbf{c}}_i)^T (\mathbf{x} - \hat{\mathbf{c}}_i) - \varepsilon^2 & \text{with } i = j - 2n \\ & \text{if } 2n+1 \leq j \leq 3n \\ \rho^2 - \mathbf{x}^T \mathbf{x} & \text{if } j = 3n+1 \end{cases} \tag{20}$$

for $j = 1, \ldots, 3n+1$, which will be used to constrain the new optimization problem. From $h(\mathbf{z})$, $l_i(\mathbf{z})$, and $m_j(\mathbf{z})$, let us define

$$p(\mathbf{z}) = q - h(\mathbf{z}) - \sum_{i=1}^{n} r_i(\mathbf{z}) l_i(\mathbf{z}) - \sum_{j=1}^{3n+1} s_j(\mathbf{z}) m_j(\mathbf{z}) \tag{21}$$

where $q \in \mathbb{R}$ is an additional variable, and $r_1(\mathbf{z}), \ldots, r_n(\mathbf{z})$ and $s_1(\mathbf{z}), \ldots, s_{3n+1}(\mathbf{z})$ are additional polynomial variables.

Theorem 1: Let us define the optimization problem

$$q^* = \min_{q, \mathbf{r}_i, \mathbf{s}_j} q \quad \text{s.t.} \quad \begin{cases} p(\mathbf{z}) \text{ is SOS} \\ s_j(\mathbf{z}) \text{ is SOS } \forall j = 1, \ldots, 3n+1. \end{cases} \tag{22}$$

Then,

$$q^* \geq g^*. \tag{23}$$

Proof: Suppose that the constraints in (22) are satisfied. This implies that $p(\mathbf{z})$ and $s_j(\mathbf{z})$, $j = 1, \ldots, 3n+1$, are nonnegative for all $\mathbf{z}$. Let us define the semialgebraic set

$$\mathcal{Z} = \{\mathbf{z} : l_i(\mathbf{z}) = 0, m_j(\mathbf{z}) \geq 0, i = 1, \ldots, n, j = 1, \ldots, 3n+1\}.$$

The condition $z \in \mathcal{Z}$ implies that

$$\begin{aligned} 0 &\leq p(\mathbf{z}) \\ &= q - h(\mathbf{z}) - \sum_{i=1}^{n} r_i(\mathbf{z}) l_i(\mathbf{z}) - \sum_{j=1}^{3n+1} s_j(\mathbf{z}) m_j(\mathbf{z}) \\ &\leq q - h(\mathbf{z}). \end{aligned}$$

Hence,

$$\mathbf{z} \in \mathcal{Z} \Rightarrow q \geq h(\mathbf{z}).$$

The condition $z \in \mathcal{Z}$ also implies that

$$y_i \left\| \hat{\mathbf{O}}_i^T (\mathbf{x} - \hat{\mathbf{c}}_i) \right\| = 1$$

since $l_i(\mathbf{z}) = 0$ and $y_i \geq 0$ from $m_j(\mathbf{z}) \geq 0$ for $j = 1, \ldots, n$. This means that

$$\begin{aligned} h(\mathbf{z}) &= \frac{1}{n} \sum_{i=1,\ldots,n} y_i \left(\hat{\mathbf{O}}_i^T (\mathbf{x} - \hat{\mathbf{c}}_i) \right)^T \hat{\mathbf{a}}_i \\ &= \frac{1}{n} \sum_{i=1,\ldots,n} \left(\frac{\hat{\mathbf{O}}_i^T (\mathbf{x} - \hat{\mathbf{c}}_i)}{\left\| \hat{\mathbf{O}}_i^T (\mathbf{x} - \hat{\mathbf{c}}_i) \right\|} \right)^T \hat{\mathbf{a}}_i \\ &= g \end{aligned}$$

and, hence,

$$\mathbf{z} \in \mathcal{Z} \Rightarrow h(\mathbf{z}) = g(\mathbf{x}).$$

Lastly, let us observe that the constraints in (16) coincide with $m_j(\mathbf{z}) \geq 0$ for $j = n+1, \ldots, 3n+1$. Summarizing, the fulfillment of the constraints in (22) implies that

$$\begin{cases} q \geq g(\mathbf{x}) \\ \mathbf{e}_3^T \hat{\mathbf{O}}_i^T (\mathbf{x} - \hat{\mathbf{c}}_i) \geq 0 \; \forall i = 1, \ldots, n \\ \|\mathbf{x} - \hat{\mathbf{c}}_i\| \geq \varepsilon \; \forall i = 1, \ldots, n \\ \rho \geq \|\mathbf{x}\| \end{cases}$$

i.e., q is an upper bound of $g(\mathbf{x})$ for all admissible values of $\mathbf{x}$. Therefore, (23) holds. □

Theorem 1 provides an upper bound of g^* through the solution of the optimization problem (22), which consists of minimizing a linear cost function subject to the condition that some polynomials are SOS. Since these polynomials depend linearly on the variables of the optimization problem, and since establishing whether a polynomial is SOS amounts to checking feasibility of an LMI, (22) is a convex optimization problem, in particular a semidefinite program. The optimization problem (22) has been built exploiting the Positivstellensatz, which allows one to investigate the nonnegativity of a polynomial over a semialgebraic set through SOS polynomials. See, e.g., [8] and references therein for details.

Solving the optimization problem (22) provides one with the upper bound q^* of g^*. In order to establish whether this upper bound is exact, and also obtain a candidate solution of the triangulation problem (8), let us proceed as follows.

Let $p^*(\mathbf{z})$ be the polynomial $p(\mathbf{z})$ evaluated for the optimal values of the variables of (22), and let $\mathbf{P}^*$ be its Gram matrix. Let $\mathbf{t}(\mathbf{z})$ denote the vector of monomials used to define $\mathbf{P}^*$, i.e.,

$$p^*(\mathbf{z}) = \mathbf{t}(\mathbf{z})^T \mathbf{P}^* \mathbf{t}(\mathbf{z}). \tag{24}$$

Let us denote the $3 \times n + 3$ matrix that extracts $\mathbf{x}$ from $\mathbf{z}$ as $\mathbf{E}$, i.e.,

$$\mathbf{x} = \mathbf{E}\mathbf{z}. \tag{25}$$

Theorem 2: The upper bound q^* is exact, i.e.,

$$q^* = g^* \tag{26}$$

if and only if there exists $\bar{\mathbf{x}}$ such that

$$\begin{cases} g(\bar{\mathbf{x}}) = q^* \\ \mathbf{e}_3^T \hat{\mathbf{O}}_i^T (\bar{\mathbf{x}} - \hat{\mathbf{c}}_i) \geq 0 \; \forall i = 1, \ldots, n \\ \|\bar{\mathbf{x}} - \hat{\mathbf{c}}_i\| \geq \varepsilon \; \forall i = 1, \ldots, n \\ \rho \geq \|\bar{\mathbf{x}}\|. \end{cases} \tag{27}$$

In such a case, $\bar{\mathbf{x}}$ is a solution of the triangulation problem (8), and there exists $\bar{\mathbf{z}}$ such that

$$\begin{cases} \mathbf{t}(\bar{\mathbf{z}}) \in \ker(\mathbf{P}^*) \\ \bar{\mathbf{x}} = \mathbf{E}\bar{\mathbf{z}}. \end{cases} \tag{28}$$

Proof: "⇐" Suppose that (27) holds. This means that there exists a scene point $\bar{\mathbf{x}}$ satisfying the constraints in (8) where the upper bound q^* is achieved. Hence, (26) holds.

"⇒" Suppose that (26) holds. Let $\mathbf{z}^*$ be the value of $\mathbf{z}$ for $\mathbf{x}$ replaced by $\mathbf{x}^*$ and $y_1, \ldots, y_n$ defined by

$$y_i \left\| \hat{\mathbf{O}}_i^T (\mathbf{x}^* - \hat{\mathbf{c}}_i) \right\| = 1.$$

Since

$$q^* = g^* = g(\mathbf{x}^*)$$

it follows that

$$q^* = h(\mathbf{z}^*).$$

Let $r_i^*(\mathbf{z})$ and $s_j^*(\mathbf{z})$ be the polynomials $r_i(\mathbf{z})$ and $s_j(\mathbf{z})$ evaluated for the optimal values of the variables of (22). It follows that

$$\begin{aligned} 0 &\leq p^*(\mathbf{z}^*) \\ &= q^* - h(\mathbf{z}^*) - \sum_{i=1}^{n} r_i^*(\mathbf{z}^*) l_i(\mathbf{z}^*) - \sum_{j=1}^{3n+1} s_j^*(\mathbf{z}^*) m_j(\mathbf{z}^*) \\ &= - \sum_{j=1}^{3n+1} s_j^*(\mathbf{z}^*) m_j(\mathbf{z}^*) \end{aligned}$$

since $l_i(\mathbf{z}^*) = 0$. This implies that

$$\sum_{j=1}^{3n+1} s_j^*(\mathbf{z}^*) m_j(\mathbf{z}^*) = 0$$

since $s_j^*(\mathbf{z}^*) \geq 0$ and $m_j(\mathbf{z}^*) \geq 0$. Therefore,

$$\begin{aligned} 0 &= p^*(\mathbf{z}^*) \\ &= \mathbf{t}(\mathbf{z}^*)^T \mathbf{P}^* \mathbf{t}(\mathbf{z}^*) \end{aligned}$$

and hence (28) holds with $\bar{\mathbf{z}}$ replaced by $\mathbf{z}^*$ since $\mathbf{P}^*$ is guaranteed to be positive semidefinite from the first constraint in (22). This also implies that (27) holds with $\bar{\mathbf{x}} = \mathbf{x}^*$. □

Theorem 2 provides a sufficient and necessary condition for establishing whether the upper bound q^* is exact, and provides a candidate solution of the triangulation problem (8). In particular, one searches for the vectors $\bar{\mathbf{z}}$ that satisfy the first constraint in (28), for instance using the technique in [9, Ch. 1, Sec. V]. From the found $\bar{\mathbf{z}}$, one extracts $\mathbf{x}$ according to the second constraint in (28) and checks whether (27) holds: if yes, then the upper bound q^* is exact and $\bar{\mathbf{x}}$ is the sought solution of (8). In all the examples presented in the next section, $\bar{\mathbf{z}}$ is simply extracted from the eigenvector corresponding to the smallest eigenvalue of $\mathbf{P}^*$.

It is worth mentioning that the exactness condition of Theorem 2 is alternative to existing exactness conditions based on rank tests for optimization over polynomials with SOS polynomials or moments, see, e.g., [8], [10], [11], and references therein. In particular, the exactness condition of Theorem 2 is based on the computation of the optimizer, which is then used to provide the candidate solution $\bar{\mathbf{x}}$ of the triangulation problem (8).

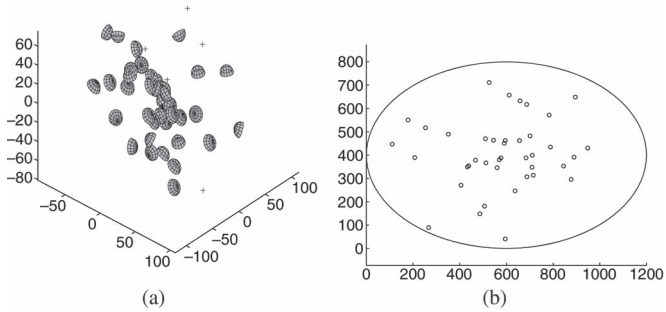


Fig. 1. Synthetic data: (a) scene points ("+" marks) and generalized cameras for 10 of the 500 vision systems; (b) image projections of such scene points ("o" marks) and boundary of the visible region (solid line).

TABLE I
SYNTHETIC DATA: AVERAGE 3-D ERROR FOR DIFFERENT LEVELS η_{IN} OF THE IMAGE NOISE

$n=2$	$\eta_{IN}=0.5$	1	1.5	2
this	4.015	3.643	3.835	4.349
algebraic	4.164	3.712	4.075	4.306
$n=3$	$\eta_{IN}=0.5$	1	1.5	2
this	1.932	1.656	1.746	1.802
algebraic	2.282	1.844	2.108	2.304
$n=4$	$\eta_{IN}=0.5$	1	1.5	2
this	1.431	1.260	1.357	1.528
algebraic	1.675	1.577	1.549	1.736

IV. EXAMPLES

In this section, we present some results obtained with synthetic and real data. The optimization problem (22) is solved with the toolbox SeDuMi [12]. The degrees of the polynomial multipliers $\mathbf{r}_i(\mathbf{z})$ and $\mathbf{s}_j(\mathbf{z})$ are chosen so that $p(\mathbf{z})$ has the lowest degree. The exactness condition of Theorem 2 is (numerically) satisfied in all the examples.

A. Synthetic Data

We have generated 500 vision systems, each of them composed by a scene point to reconstruct (denoted hereafter as $\mathbf{x}$) and four generalized cameras with 180 degrees field of view, in particular with parameter $\xi = 0.5$ and intrinsic parameters given by

$$\mathbf{K} = \begin{pmatrix} 300 & 0 & 600 \\ 0 & 200 & 400 \\ 0 & 0 & 1 \end{pmatrix}.$$

For each vision system, $\mathbf{x}$ and the centers of the cameras are randomly chosen in a sphere of radius 100 centered in the origin of the reference frame, while the orientation matrices of the cameras are randomly chosen under the constraint that $\mathbf{x}$ is visible by the cameras. Fig. 1 illustrates some of the generated vision systems.

In order to generate the corrupted data, we have:

- introduced image noise by adding random variables in the interval $[-\eta_{IN}, \eta_{IN}]$ pixels to each coordinate of the image points, where η_{IN} defines the level of the noise;
- introduced calibration errors by multiplying each intrinsic parameter (ξ and all the entries of $\mathbf{K}$) and each extrinsic parameter (all the entries of $\mathbf{c}_i$ and $\mathbf{O}_i$, $i = 1, \ldots, n$) times random variables in the interval $[1 - \eta_{CE}/100, 1 + \eta_{CE}/100]$, where η_{CE} defines the level of the errors.

First, we have tested the proposed method for different levels η_{IN} of the image noise. Tables I and II show the obtained average 3-D error and the average reprojection error (denoted by "this"). These tables also show the results obtained by minimizing the algebraic error (see, e.g., [1]) with standard linear least-squares (denoted by "algebraic").

TABLE II
SYNTHETIC DATA: AVERAGE REPROJECTION ERROR FOR DIFFERENT LEVELS η_{IN} OF THE IMAGE NOISE

$n=2$	$\eta_{IN}=0.5$	1	1.5	2
this	3.291	3.286	3.266	3.255
algebraic	3.802	3.616	3.610	3.693
$n=3$	$\eta_{IN}=0.5$	1	1.5	2
this	2.861	2.905	2.993	2.778
algebraic	3.450	3.417	3.568	3.488
$n=4$	$\eta_{IN}=0.5$	1	1.5	2
this	2.784	2.842	2.823	2.733
algebraic	3.014	3.027	3.083	3.190

TABLE III
SYNTHETIC DATA: AVERAGE 3-D ERROR FOR DIFFERENT LEVELS η_{CE} OF THE CALIBRATION ERRORS

$n=2$	$\eta_{CE}=0.5$	1	1.5	2
this	1.660	6.448	4.787	6.457
algebraic	1.771	6.508	5.352	6.900
$n=3$	$\eta_{CE}=0.5$	1	1.5	2
this	0.975	1.690	2.509	3.814
algebraic	1.135	2.013	2.713	4.362
$n=4$	$\eta_{CE}=0.5$	1	1.5	2
this	0.627	1.286	1.867	2.540
algebraic	0.752	1.684	2.255	3.181

TABLE IV
SYNTHETIC DATA: AVERAGE REPROJECTION ERROR FOR DIFFERENT LEVELS η_{CE} OF THE CALIBRATION ERRORS

$n=2$	$\eta_{CE}=0.5$	1	1.5	2
this	1.652	3.198	5.046	6.090
algebraic	1.928	3.674	5.696	18.738
$n=3$	$\eta_{CE}=0.5$	1	1.5	2
this	1.567	3.104	4.565	5.680
algebraic	1.900	3.703	4.815	6.859
$n=4$	$\eta_{CE}=0.5$	1	1.5	2
this	1.380	2.614	4.266	5.288
algebraic	1.576	3.022	4.459	6.138

TABLE V
SYNTHETIC DATA: AVERAGE 3-D ERROR FOR DIFFERENT LEVELS η_{IN} OF THE IMAGE NOISE IN THE PRESENCE OF OUTLIERS

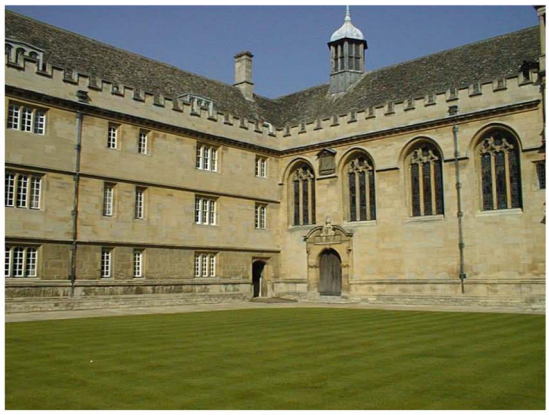
$n=4$	$\eta_{IN}=0.5$	1	1.5	2
this	2.738	4.773	6.864	9.838
algebraic	3.284	5.574	8.870	11.338

TABLE VI
SYNTHETIC DATA: AVERAGE REPROJECTION ERROR FOR DIFFERENT LEVELS η_{IN} OF THE IMAGE NOISE IN THE PRESENCE OF OUTLIERS

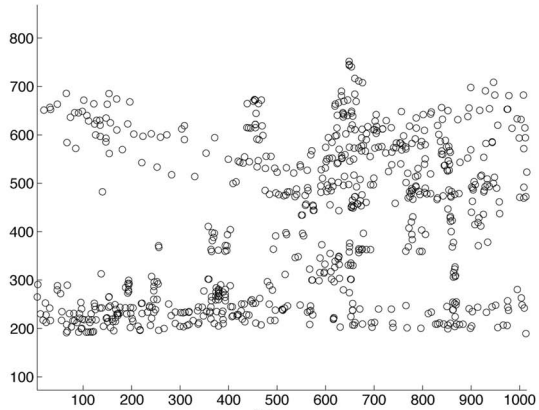
$n=4$	$\eta_{IN}=0.5$	1	1.5	2
this	5.044	9.340	14.018	16.471
algebraic	5.914	10.697	16.832	20.165

Second, we have repeated the test for different levels η_{CE} of the calibration errors, obtaining the results shown by Tables III and IV.

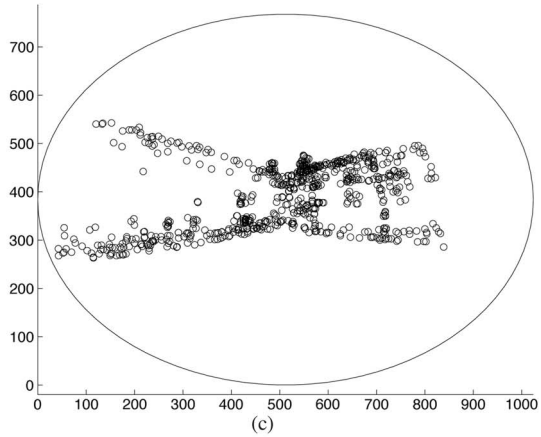
Lastly, we have investigated the behavior in the presence of outliers, i.e., a few image points incorrectly matched. To this end, we have considered the case of four generalized cameras where an outlier is present on one camera by setting the level η_{IN} of the image noise equal to 100 pixels. Tables V and VI show the obtained results.



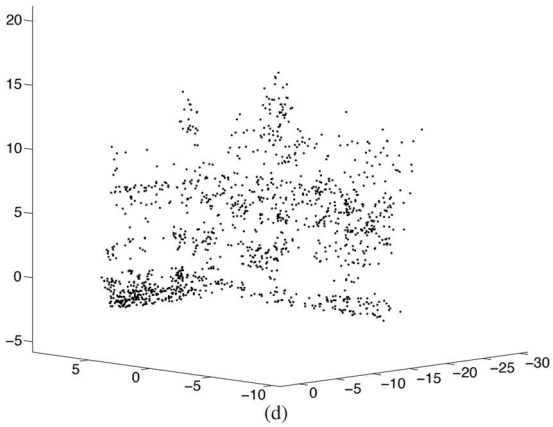
(a)



(b)

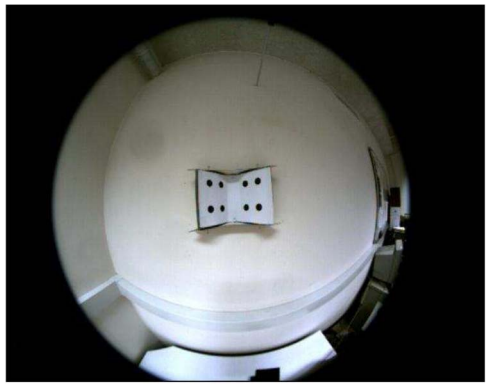


(c)



(d)

Fig. 2. Real data (Wadham college sequence): (a) one of the five images; (b) points extracted in such an image; (c) same points after transforming and shifting the cameras; (d) estimated scene points.



(a)

(b)

Fig. 3. Real data (laboratory): (a) one of the three images of the dots pattern; (b) one of the three images of the Maruko-chan doll.

B. Real Data: Wadham College Sequence

Here we consider the Wadham college sequence available at the webpage of the Visual Geometry Group of Oxford University. This sequence consists of five views taken with a perspective camera, the projection matrices of such views, and 3019 image points corresponding to 1331 scene points visible in at least two of such views (with known correspondence). Fig. 2(a) shows one of these views.

First, we have estimated the 1331 scene points using standard triangulation for perspective cameras. Second, we have computed the projections of these scene points onto generalized cameras with same orientation, same center except for a translation along the optical axis in order to enlarge the spanned image area, and parameter $\xi = 0.5$. Fig. 2(b) and (c) show this process for the points of the view in Fig. 2(a). The data obtained so far are used as true data. Third, we have corrupted the true data as done in the previous subsection for the case of synthetic data with noise intensity $\eta_{IN} = \eta_{CE} = 0.2$. Fourth, we have repeated the triangulation using for each scene point the maximum number of cameras where the point is visible.

Fig. 2(d) shows the estimated scene points. The average 3-D error obtained with the proposed method is 0.396, while the one obtained by minimizing the algebraic error is 0.446. The average reprojection error obtained with the proposed method is 0.476, while the one obtained by minimizing the algebraic error is 0.480.

C. Real Data: Laboratory

Lastly, we present the results obtained with some real data acquired in our laboratory. We have taken three images from different viewpoints of two objects, a dots pattern and a Maruko-chan doll, see Fig. 3(a) and (b) where an image is shown for each object. The problem

TABLE VII
REAL DATA (LABORATORY): REPROJECTION ERROR FOR THE OBJECT IN FIG. 3(a)

point no.	1	2	3	4	5	6	7	8
this	2.563	2.562	3.194	3.666	3.692	3.340	3.286	3.823
algebraic	2.611	2.624	3.281	3.746	3.751	3.412	3.355	3.877

TABLE VIII
REAL DATA (LABORATORY): REPROJECTION ERROR FOR THE OBJECT IN FIG. 3(b)

point no.	1	2	3	4	5	6	7
this	2.453	2.100	0.801	3.052	2.424	2.232	2.099
algebraic	2.602	2.233	0.868	3.203	2.557	2.280	2.169

consists of estimating the positions of the eight dots of the pattern, and of seven distinctive points of the doll.

The camera is coarsely calibrated, in particular the estimate for ξ is $\hat{\xi} = 0.735$ and the one for $\mathbf{K}$ is

$$\hat{\mathbf{K}} = \begin{pmatrix} 269.9 & 0 & 316.8 \\ 0 & 298.4 & 254.9 \\ 0 & 0 & 1 \end{pmatrix}.$$

Tables VII and VIII show the reprojection error obtained for each estimated point.

V. CONCLUSION

We have proposed a method for estimating the position of a scene point from its image projections onto cameras that can be modeled by a spherical projection followed by a perspective one. The sought estimate is defined by minimizing the angles between the projections on the sphere of the available image points and the corresponding projections of the estimate. A solution based on convex optimization with LMIs has been proposed for addressing this problem, which provides a candidate of the sought estimate. Moreover, a condition has been provided for establishing exactness of the found candidate, i.e., establishing whether the found candidate is a minimizer of the considered geometric criterion.

REFERENCES

[1] R. Hartley and A. Zisserman, *Multiple View in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
[2] G. Chesi and Y. S. Hung, "Global path-planning for constrained and optimal visual servoing," *IEEE Trans. Robotics*, vol. 23, no. 5, pp. 1050–1060, Oct. 2007.
[3] R. Hartley and P. Sturm, "Triangulation," *Comput. Vis. Image Understand.*, vol. 68, no. 2, pp. 146–157, 1997.
[4] H. Stewenius, F. Schaffalitzky, and D. Nister, "How hard is 3-view triangulation really?" in *Proc. Int. Conf. Comput. Vis.*, Beijing, China, 2005, pp. 686–693.
[5] G. Chesi and Y. S. Hung, "Fast multiple-view L2 triangulation with occlusion handling," *Comput. Vis. Image Understand.*, vol. 115, no. 2, pp. 211–223, 2011.
[6] G. Chesi, "A geometric approach to multiple-view triangulation for fish-eye cameras," in *Proc. IAPR Conf. Mach. Vis. Applicat.*, Kyoto, Japan, 2013, pp. 210–213.
[7] O. Tahri, Y. Mezouar, F. Chaumette, and H. Araujo, "Visual servoing and pose estimation with cameras obeying the unified model," in *Proc. Vis. Servoing via Adv. Numer. Meth.*, G. Chesi and K. Hashimotos, Eds., 2010, pp. 231–250, Springer.
[8] G. Chesi, "LMI techniques for optimization over polynomials in control: A survey," *IEEE Trans. Autom. Control*, vol. 55, no. 11, pp. 2500–2510, Nov. 2010.
[9] G. Chesi, *Domain of Attraction: Analysis and Control via SOS Programming*. New York, NY, USA: Springer, 2011.
[10] J.-B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM J. Optimiz.*, vol. 11, no. 3, pp. 796–817, 2001.
[11] P. A. Parrilo, "Semidefinite programming relaxations for semialgebraic problems," *Math. Program. Ser. B*, vol. 96, no. 2, pp. 293–320, 2003.
[12] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimix. Meth. Software*, vol. 11–12, pp. 625–653, 1999.
[13] G. Chesi, "Camera displacement via constrained minimization of the algebraic error," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 370–375, 2009.